



```

graph TD
    Input[Input: 28x28x1] --> Conv2d_1[conv2d_1: input: (None, 28, 28, 1)  
input_layer: output: (None, 28, 28, 3)]
    Conv2d_1 --> Conv2d_2[conv2d_2: input: (None, 28, 28, 3)  
Conv2D: output: (None, 28, 28, 32)]
    Conv2d_2 --> Activation_1[activation_1: input: (None, 28, 28, 32)  
Activation: output: (None, 28, 28, 32)]
    Activation_1 --> Batch_Norm_1[batch_normalization: input: (None, 28, 28, 32)  
BatchNormalization: output: (None, 28, 28, 32)]
    Batch_Norm_1 --> Max_Pooling_1[max_pooling2d: input: (None, 28, 28, 32)  
MaxPooling2D: output: (None, 14, 14, 32)]
    Max_Pooling_1 --> Dropout_1[dropout: input: (None, 14, 14, 32)  
Dropout: output: (None, 14, 14, 32)]
    Dropout_1 --> Conv2d_3[conv2d_3: input: (None, 14, 14, 32)  
Conv2D: output: (None, 14, 14, 32)]
    Conv2d_3 --> Activation_2[activation_2: input: (None, 14, 14, 32)  
Activation: output: (None, 14, 14, 32)]
    Activation_2 --> Batch_Norm_2[batch_normalization_2: input: (None, 14, 14, 32)  
BatchNormalization: output: (None, 14, 14, 32)]
    Batch_Norm_2 --> Max_Pooling_2[max_pooling2d_2: input: (None, 14, 14, 32)  
MaxPooling2D: output: (None, 7, 7, 32)]
    Max_Pooling_2 --> Dropout_2[dropout_2: input: (None, 7, 7, 32)  
Dropout: output: (None, 7, 7, 32)]
    Dropout_2 --> Conv2d_4[conv2d_4: input: (None, 7, 7, 32)  
Conv2D: output: (None, 7, 7, 10)]
    Conv2d_4 --> Activation_3[activation_3: input: (None, 7, 7, 10)  
Activation: output: (None, 7, 7, 10)]
    Activation_3 --> Batch_Norm_3[batch_normalization_3: input: (None, 7, 7, 10)  
BatchNormalization: output: (None, 7, 7, 10)]
    Batch_Norm_3 --> Dense_1[dense_1: input: (None, 7, 7, 10)  
Dense: output: (None, 1280)]
    Dense_1 --> Dense_2[dense_2: input: (None, 1280)  
Dense: output: (None, 1024)]
    Dense_2 --> Dense_3[dense_3: input: (None, 1024)  
Dense: output: (None, 10)]
    Dense_3 --> Output[Output: 10]
  
```

The graph displays the training and validation loss over 50 epochs. The training loss (blue line) shows a consistent downward trend, starting at approximately 2.5 and ending near 1.0. The validation loss (orange line) starts at a high value of about 4.8, drops sharply to around 1.5 by epoch 15, and then exhibits significant fluctuations, generally staying between 1.2 and 2.5. This pattern suggests that the model is learning to fit the training data well but may be overfitting after epoch 15, as the validation loss does not continue to decrease and instead fluctuates at a higher level than the training loss.

